

User manual

**Strategies for grouping
chemicals to fill data gaps to
assess genetic toxicity and
genotoxic carcinogenicity**

Document history

Version	Comment
Version 1.0	April 2010, first version
Version 1.1	July 2013, second version, updated for version 3.1 of the Toolbox

Issue date: July 2013

Language: English

If you have questions or comments that relate to this document, please send them to ehscont@oecd.org or visit the QSAR Toolbox discussion forum at https://community.oecd.org/community/toolbox_forum

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

Table of Contents

1. How to use this guidance document	5
2. Genetic toxicity and carcinogenicity endpoints.....	6
2.1 In vitro endpoints.....	6
2.2 In vivo endpoints.....	7
3. Primary profilers relevant to genetic toxicity and carcinogenicity endpoints ...	7
3.1 DNA binding by OECD	10
3.2 DNA binding by OASIS v1.1	10
3.3 Carcinogenicity (genotoxicity and non-genotoxicity) by ISS	11
3.4 In vitro mutagenicity (Ames test) by ISS	11
3.5 In vivo mutagenicity (micronucleus) by ISS.....	11
3.6 OncoLogic primary classification	12
4. Databases relevant to genetic toxicology and carcinogenicity.....	13
5. Profiling results: What they say about a grouping strategy.....	14
5.1 Ideal profiling scenario.....	14
5.1.1 Ethanal – mutagenicity	15
5.1.2 Ethylene oxide – chromosomal mutation.....	15
5.2 Reasonable profiling scenario.....	16
5.2.1 N-nitrosodimethylamine – genotoxic carcinogenicity	16
5.3 Unreliable profiling scenario.....	17
5.3.1 Trichlorotoluene - genotoxic carcinogenicity	17
5.4 General conclusions regarding the outcome of profiling strategy results	18
5.4.1 The use of multiple mechanistic profilers	19
5.4.2 The use of endpoint specific profilers	19
6. Secondary profilers relevant to genetic toxicology and genotoxic carcinogenicity	20
6.1 Defining the structural domain of a chemical category.....	21
7. Initial category formation using the primary profilers to define the mechanistic domain.....	23
7.1 Mutagenicity	24
7.2 Chromosomal mutation	26
7.3 Genotoxic carcinogenicity.....	28
7.4 General conclusions regarding the sub-categorisation with the primary profilers.....	30
8. Definition of the structural domain and data-gap filling using read across.....	31
8.1 Structural domain definition using the secondary profilers	31

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

8.2 Data-gap filling via read across 31

9. General approach for the development of chemical categories for genotoxicity..... 33

10. References 35

Table of Tables

Table 3.1: Primary profilers for genetic toxicity and carcinogenicity available in the OECD QSAR Toolbox 9

Table 4.1: Summary of databases available within the OECD QSAR Toolbox that are relevant to genetic toxicity and carcinogenicity.....13

Table 5.1: Summary of the applicable primary profilers for the mutagenicity, chromosomal mutation and genotoxic carcinogenicity available in the OECD QSAR Toolbox19

Table 6.1: Secondary profilers relevant to genetic toxicity and genotoxic carcinogenicity endpoints available in the OECD QSAR Toolbox20

Table of Figures

Figure 5.1: Result of primary profiling for ethanal using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox.15

Figure 5.2: Result of primary profiling for ethylene oxide using a combination of the mechanistic and endpoint specific profilers for available for chromosomal mutation in the OECD QSAR Toolbox.16

Figure 5.3: Result of primary profiling for *N*-nitrosodimethylamine using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox.17

Figure 5.4: Result of primary profiling for trichlorotoluene using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox.....18

Figure 7.1: Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler25

Figure 7.2: Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler27

Figure 7.3: Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OASIS' profiler29

Figure 8.1: Read across prediction made for the Ames mutagenicity for ethanal in the TA 100 strain of *Salmonella* in the presence of the S9 liver fraction.32

Figure 9.1: General scheme for category formation for genetic toxicity and genotoxic carcinogenicity34

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

1 How to use this guidance document

This document is intended to offer guidance on how to form robust chemical categories suitable for data gap filling for genetic toxicity and genotoxic carcinogenicity in the OECD QSAR Toolbox. The endpoints covered in this guidance document all feature covalent DNA binding as the molecular initiating event. The examples used in this document are intended to illustrate one approach to the formation of chemical categories with the OECD QSAR Toolbox. It is important to realise that no effort has been made to validate the read across predictions (by analysing data sources not available in V3 of the OECD QSAR Toolbox). The term genetic toxicity and genotoxic carcinogenicity will refer to the following endpoints in this document:

- Gene mutation (Ames mutagenicity assay)
- *In vitro* and *in vivo* chromosomal mutation (micronucleus assay)
- Genotoxic carcinogenicity

The document is split into several sections, these being:

- Sections 2 – 4: Introductory material about the profilers and databases available for genetic toxicity and genotoxic carcinogenicity. These sections include additional references to the literature that provide more in depth background material.
- Sections 5 – 8: Worked examples for profiling of target chemicals and how to use this information to form chemical categories for each of the endpoints. These sections are intended to be used as examples that the reader can follow as illustrations of some of the recommended strategies. A first time reader of this document should spend time ensuring that they can use the profilers and form the suggested categories as instructed. A good working knowledge of the OECD QSAR Toolbox is required for these examples to be of benefit and thus users who do not consider themselves as experts should consult the appropriate guidance documents available from the OECD [1].
- Section 9: Summary of a general strategy that can be used to generate chemical categories suitable for read across for a user's own data. This summary was used to generate the example categories covered in this guidance document. The worked examples in sections 5 – 8 should be undertaken before attempting to use this summary information.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

2 Genetic toxicity and carcinogenicity endpoints

This guidance document aims to outline strategies for grouping chemicals into chemical categories for genetic toxicity and genotoxic carcinogenicity. The effects covered in this document are:

- *In vitro* gene mutation (Ames mutagenicity)
- *In vitro* chromosomal mutation (micronucleus assay)
- *In vivo* chromosomal mutation (micronucleus assay)
- *In vivo* genotoxic carcinogenicity (rodent assay)

These endpoints are assumed to share a common molecular initiating event [2] in the covalent binding of a chemical to DNA [3-8]. Other molecular initiating events exist, especially for carcinogenicity, such as protein binding, non-covalent interactions with protein receptors, intercalation with DNA and the formation of free radicals (for details of these mechanisms see reference [8]). The focus of this guidance document is solely genotoxicity.

2.1 *In vitro* endpoints

This guidance document covers the two *in vitro* effects that are typically required for the regulatory assessment of a chemical's mutagenicity. The first of these is the Ames mutagenicity assay (OECD guideline 471). This is specifically designed to assess the ability of a chemical to cause point mutations in the DNA of the bacterium *Salmonella typhimurium* [9, 10]. A number of bacterial strains (TA1537, TA1535, TA100, TA98 and TA97) have been engineered to detect differing classes of mutagenic chemicals. The basic test only detects directly acting mutagens (those chemicals able to interact with DNA without the need for metabolic activation). The inclusion of an S9 mix of rodent liver enzymes is designed to assess those chemicals requiring bioactivation in order to be mutagenic. Typically, chemicals are assayed with and without the presence of the S9 mix. The assay results are reported in a binary fashion, with a positive result in any of the bacterial strains confirming mutagenic potential.

The second *in vitro* test discussed in this document is the micronucleus assay for chromosomal mutation (OECD guideline 487). The *in vitro* micronucleus assay involves the use of rodent or human cell lines to detect chemicals that can bind to chromosomes resulting in the production of micronuclei (small membrane bound fragments of DNA) during nuclear division. The production of micronuclei is an indicator of damage to either the chromosome and/or the ability of the cell to divide correctly [11]. As with the Ames assay, the micronucleus test is conducted with and without the inclusion of the S9 mix of rodent liver enzymes. The *in vitro* micronucleus assay results are also reported in a binary fashion, with a positive result indicating genotoxic potential.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

2.2 In vivo endpoints

In addition to the two *in vitro* tests described in section 1.1, the use of data from two *in vivo* tests will also be discussed in this guidance document. The first is the *in vivo* micronucleus assay carried out in rodents (OECD guideline 474). As with the *in vitro* micronucleus assay, this assay is specifically designed to detect chemicals capable of causing the production of micronuclei during nuclear division. The *in vivo* assay has several advantages over the *in vitro* micronucleus assay in that metabolism, pharmacokinetics and potential DNA repair processes are all taken into account. A positive result is confirmed by the presence of micronuclei in either extracted bone marrow or blood samples from the animal. This indicates that the chemical is able to produce micronuclei (under the test conditions) in rodents and, thus, potentially genotoxic in humans. A negative result indicates (under the test conditions) that in rodents the chemical is not able to produce micronuclei.

The OECD QSAR Toolbox contains data from a number of experimental protocols relating to the assessment of carcinogenicity (OECD guideline 451). In general the assessment of carcinogenicity involves administering a chemical orally to rodents and recording its ability to cause cancerous lesions in a variety of tissues. Two common results are reported for carcinogenicity studies; namely histopathology and TD₅₀. The results of histopathology examinations are reported as a positive/negative, with a positive result indicating the presence of a tumour in at least one of the tissues examined. The second result, TD₅₀ is the standardised potency measured as the daily dose-rate in mg/kg/body weight/day for life to induce tumours in half of the test animals that would have remained tumour-free at zero dose [12]. Importantly, in addition to covering metabolism, pharmacokinetics and potential DNA repair mechanisms, carcinogenicity assays also cover both genotoxic and non-genotoxic mechanisms of action.

3 Primary profilers relevant to genetic toxicity and carcinogenicity endpoints

The primary profilers relevant to genetic toxicity and carcinogenicity endpoints can be divided into two types. The mechanistic primary profilers have been developed from knowledge related to the underlying chemistry governing the ability of a chemical to covalently bind to DNA. These profilers define fragments (so-called structural alerts) within chemicals that have been shown to be associated with a given reaction mechanism known to be important in covalent binding to DNA. Importantly, there are not necessarily toxicological data associated with these structural alerts – with a number of the structural alerts being derived from the results of *in chemico* testing (defined as any data taken from chemistry-based studies [13]). The second type of primary profilers have been developed using toxicological data for a given assay. These profilers define a series of structural alerts that are associated with toxicity for a given assay.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

The available profilers in the OECD QSAR Toolbox, along with how they have been developed, are summarised in Table 3.1.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

Table 3.1: Primary profilers for genetic toxicity and carcinogenicity available in the OECD QSAR Toolbox

Profiler name	Type	Applicable endpoint	Num. Alerts	Data sources that the structural alerts have been derived from	Section
DNA binding by OECD	Mechanistic	Covalent DNA binding	86	Ames Carcinogenicity Idiosyncratic drug toxicity <i>In chemico</i>	3.1
DNA binding by OASIS v1.1	Mechanistic	Covalent DNA binding	82	Ames	3.2
Carcinogenicity (genotoxic and non-genotoxic) by ISS	Endpoint	Genotoxic and non-genotoxic carcinogenicity	37	Carcinogenicity	3.3
<i>In vitro</i> mutagenicity (Ames test) by ISS	Endpoint	<i>In vitro</i> mutagenicity	30	Ames	3.4
<i>In vivo</i> mutagenicity (micronucleus) by ISS	Endpoint	<i>In vivo</i> chromosomal mutation	35	Micronucleus	3.5
OncoLogic primary classification	Endpoint	Carcinogenicity	48	Carcinogenicity	3.6

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

3.1 DNA binding by OECD

The "DNA binding by OECD" profiler is based solely on structural alerts for the electrophilic reaction chemistry associated with covalent DNA binding. A detailed review of the electrophilic reaction chemistry covered by this profiler is given by Enoch and Cronin [5]. In the simplest terms, the applicable electrophilic reaction chemistry can be defined as the formation of a new chemical bond between a DNA base containing a nucleophilic centre [an area of the molecule with a (partial) negative charge, typically a lone pair of electrons on a nitrogen or oxygen atom] and an exogenous chemical containing an electrophilic centre [an area of a molecule with a (partial) positive charge]. Thus, the profiler details a range of structural alerts that contain electrophilic centres or those that can be metabolically activated to electrophiles. Importantly, the data used to develop this profiler were not necessarily from regulatory toxicological endpoints (such as the Ames test). Instead a range of data sources covering experiments in which a chemical had been shown to bind covalently to one of the nucleobases of DNA. For example, mutagenicity and carcinogenicity data, *in chemico* data and data from idiosyncratic drug toxicity studies were considered to develop the structural alerts. The structural alerts were assigned to one of six mechanistic domains based on the definitions developed by Aptula and Roberts [13]. These domains are acylation, unimolecular nucleophilic substitution (S_N1), bimolecular nucleophilic substitution (S_N2), Michael addition, Schiff base formation and radical formation.

3.2 DNA binding by OASIS v1.1

This profiler is a mechanistic profiler developed from an analysis of Ames mutagenicity data. The structural alerts within this profiler are detailed by Mekenyan et al and Serafimova et al [14, 15]. It contains a number of structural alerts that have been shown to be related to established electrophilic reaction chemistry known to be important in covalent DNA binding. The chemical categories developed using this profiler can be considered similar to those obtained at the structural alert level using the 'DNA binding by OECD' profiler. The electrophilic reaction chemistry for each structural alerting group is detailed within the profiler's metadata. In addition, the metadata included in the OECD QSAR Toolbox also detail examples of DNA adducts for each structural alerting group with references to toxicological data sources. This is useful information that can help to support a mechanistic analysis.

Clearly there is an overlap between the 'DNA binding by OASIS' profiler and the 'DNA binding by OECD' profiler as Ames data have been used in the development of both profilers. It is important to realise that the 'DNA binding by OASIS' profiler is focussed on well-established electrophilic mechanisms. In contrast to the 'DNA

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

binding by OECD' profiler, only structural alerts that have supporting toxicological data are included in the profiler.

3.3 Carcinogenicity (genotoxicity and non-genotoxicity) by ISS

This profiler is an expanded and updated version of the correspondent module of the software Toxtree [16]. It works as a decision tree for estimating carcinogenicity, based on a list of 55 structural alerts. Out of them, 35 derive from the Toxtree module and 20 are newly derived. Most of the new structural alerts are relative to non-genotoxic carcinogenicity, whereas the structural alerts in the initial list mainly coded genotoxic carcinogenicity. The structural alerts for carcinogenicity are molecular functional groups or substructures known to be linked to the carcinogenic activity of chemicals. As one or more structural alerts embedded in a molecular structure are recognised, the system flags the potential carcinogenicity of the chemical [17].

3.4 In vitro mutagenicity (Ames test) by ISS

This profiler is based on the Mutagenicity/Carcinogenicity module of the Toxtree software. It works as a decision tree for estimating *in vitro* (Ames test) mutagenicity, based on a list of 30 structural alerts. The structural alerts for mutagenicity are molecular functional groups or substructures known to be linked to the mutagenic activity of chemicals. As one or more structural alerts embedded in a molecular structure are recognised, the system flags the potential mutagenicity of the chemical. The present list of structural alerts is a subset of the original Toxtree list, obtained by eliminating the structural alerts for non-genotoxic carcinogenicity [17].

3.5 In vivo mutagenicity (micronucleus) by ISS

This profiler consists of a list of 35 structural alerts developed from a preliminary screening of potential *in vivo* mutagens. These structural alerts are molecular functional groups or substructures that are known to be linked to the induction of effects in the *in vivo* micronucleus assay. The compilation of structural alerts for the *in vivo* micronucleus assay in rodents provided here is based on both the existing knowledge on the mechanisms of toxic action and on a structural analysis of the chemicals tested in the assay.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

3.6 OncoLogic primary classification

The OncoLogic primary classification profiler was developed to be used specifically to address carcinogenicity potential. The profiler was developed by the Laboratory of Mathematical Chemistry (LMC) solely to mimic the structural classes of known/potential carcinogens covered in version 7.0 of the United States Environmental Protection Agency's (US EPA) OncoLogic Cancer Expert System for predicting carcinogenic potential [18]. No attempts were made to incorporate the additional expert system rule base from the OncoLogic software into the OECD QSAR Toolbox. Since the additional enhancing/mitigating rules of OncoLogic are not part of the OECD QSAR Toolbox profiler, users should note that classification of a query chemical in an OncoLogic class does not automatically mean that the chemical will be a carcinogen. As with the other endpoint specific profilers, this means that for each structural alerting group within the profiler, there is at least a single reported incidence of a chemical causing carcinogenicity. However, the same caveat also applies in that the presence of a structural alerting group within a query chemical does not necessarily indicate that the query chemical will be a carcinogen. In addition, no supporting mechanistic chemistry is available for this profiler (within the OECD QSAR Toolbox).

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

4 Databases relevant to genetic toxicology and carcinogenicity

The OECD QSAR Toolbox contains a number of databases that are relevant to genetic toxicology and carcinogenicity. These have been donated by various organisations and cover a range of regulatory endpoints and species. These databases are summarised as shown in Table 4.1.

Table 4.1: Summary of databases available within the OECD QSAR Toolbox that are relevant to genetic toxicity and carcinogenicity

Database	Num. Chemicals	Endpoint	Species
Bacterial mutagenicity ISSSTY	7367	Gene mutation	<i>In vitro</i> : <i>Salmonella typhimurium</i>
Carcinogenic potency database CPDB	1530	Carcinogenicity (TD ₅₀)	<i>In vivo</i> : Rats, mice, hamsters, dogs and non human primates
Carcinogenicity ISSCAN	1150	Carcinogenicity (TD ₅₀)	<i>In vivo</i> : Rats and mice
Cell Transformation Assay ISSCTA	327	Non-genotoxic carcinogenicity	<i>In vitro</i> : Syrian Hamster Embryo (SHE) cells BALB/c 3T3 C3H/10HT1/2 Bhas42
Genotoxicity OASIS	7500	Gene mutation (yes/no) Chromosomal mutation (yes/no)	<i>In vitro</i> : <i>Salmonella typhimurium</i> <i>In vitro</i> : Chinese hamster lung cells <i>In vitro</i> : T-lymphoma cell lines
Micronucleus ISSMIC	564	Chromosomal mutation (yes/no)	<i>In vivo</i> : Rats and mice
Micronucleus OASIS	557	Chromosomal mutation	<i>In vivo</i>

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

5 Profiling results: What they say about a grouping strategy

The QSAR Toolbox aims to group chemicals into categories on the basis of a common molecular initiating event. This allows data-gaps to be filled via trend analysis and/or read across. As discussed, one common molecular initiating event for genetic toxicity and genotoxic carcinogenicity is the ability of a chemical to bind covalently to DNA. Thus, when using the QSAR Toolbox to group chemicals into categories, the following general steps for a target chemical for which a data-gap exists would be undertaken.

1. Profile the target chemical for potential mechanism of action related to the molecular initiating event
2. Use the result of this mechanistic profiling to select chemical analogues from endpoint specific toxicological databases
3. Define the mechanistic and structural domain of the resulting chemical category
4. Fill data-gaps using trend analysis and/or read across

The following sections outline how one should interpret the common profiling results for target chemicals that have covalent DNA binding as the molecular initiating event:

5.1 Ideal profiling scenario

The ideal profiling scenario occurs when one (or both) of the mechanistic profilers identifies a single mechanism related to covalent DNA binding that is supported by a single structural alert identified by the appropriate endpoint specific profiler/profilers. Importantly, it does not matter whether only one of the mechanistic profilers or both identify the single mechanism (as long as if both do they identify the same mechanism). This is because there is significant overlap between the structural alerts contained within these profilers as they both outline the chemistry associated with covalent DNA binding. Confidence in the profiling results is gained by the appropriate endpoint specific profiler identifying a complementary structural alert. The confidence is gained due to the fact that the endpoint specific profilers contain only structural alerts that have toxicological data associated with them. Consider the following examples:

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

5.1.1 Ethanal – mutagenicity

Profiling ethanal for mechanisms related to Ames mutagenicity shows that the ‘DNA binding by OECD’ profiler identifies a potential electrophilic mechanism, Schiff base formation. In contrast, no alert is triggered by ethanol in the ‘DNA binding by OASIS v1.1’ profiler. This indicates that this profiler contains no information about this type of chemical (the chemical is out of the domain of this profiler). However, the mechanistic profiling results from the ‘DNA binding by OECD’ profiler’ is supported by the results of the endpoint profiling that shows there to be an alert supported by toxicological data for simple aldehydes (Figure 5.1). These are ideal profiling results in that one of the mechanistic profilers has identified a clear mechanism that is supported by a structural alert in the appropriate endpoint specific profiler.

1 [target]	
Structure	
<input checked="" type="checkbox"/> Substance Identity	
<input checked="" type="checkbox"/> Physical Chemical Properties	
<input checked="" type="checkbox"/> Environmental Fate and Transport	
<input checked="" type="checkbox"/> Ecotoxicological Information	
<input checked="" type="checkbox"/> Human Health Hazards	
<input type="checkbox"/> Profile	
<input type="checkbox"/> General Mechanistic	
DNA binding by OASIS v.1.1	
DNA binding by OECD	
<input type="checkbox"/> Endpoint Specific	
in vitro mutagenicity (Ames test) alerts by ISS	Simple aldehyde
	No alert found
	Schiff base formers
	Schiff base formers >> Direct Acting Schiff Base Formers
	Schiff base formers >> Direct Acting Schiff Base Formers >> Mono aldehydes

Figure 5.1: Result of primary profiling for ethanal using a combination of the mechanistic and endpoint specific profilers for available for mutagenicity in the OECD QSAR Toolbox.

5.1.2 Ethylene oxide – chromosomal mutation

A similar profiling analysis can be carried out for chromosomal mutation. The profiling results show that both mechanistic profilers identify the same mechanism; an S_N2 ring opening reaction. The mechanistic profiling is supported by the appropriate endpoint specific profiler for the micronucleus endpoint (Figure 5.2). However, it is important to realise that the confidence in this category is the same as for the previous example (ethanal) with no additional weight of evidence being gained from the two mechanistic profilers being in agreement.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

1 [target]	
Structure	
Substance Identity	
Physical Chemical Properties	
Environmental Fate and Transport	
Ecotoxicological Information	
Human Health Hazards	
Profile	
General Mechanistic	
DNA binding by OASIS v.1.1	SN2 SN2 >> Direct Acting Epoxides and Related SN2 >> Direct Acting Epoxides and Related >> Epoxides, Aziridines
DNA binding by OECD	SN2 SN2 >> Direct Acting Epoxides and related SN2 >> Direct Acting Epoxides and related >> Epoxides
Endpoint Specific	
in vivo mutagenicity (Micronucleus) alerts by ISS	Epoxides and aziridines

Figure 5.2: Result of primary profiling for ethylene oxide using a combination of the mechanistic and endpoint specific profilers for available for chromosomal mutation in the OECD QSAR Toolbox.

5.2 Reasonable profiling scenario

There is the potential for one (or both) of the mechanistic profilers to identify multiple potential mechanisms related to covalent DNA binding. Such profiling results are clearly not ideal; however, if an endpoint specific profiler identifies a single structural alert that complements the mechanistic profilers then the profiling results are still useful. In cases such as this it is important to inspect the results of the mechanistic profiling and the endpoint specific profiling to ensure that it is a single structural feature that is being identified. Consider the following example:

5.2.1 N-nitrosodimethylamine – genotoxic carcinogenicity

Profiling *N*-nitrosodimethylamine for genotoxic carcinogenicity shows the two mechanistic profilers to identify multiple potential mechanisms. It is difficult from these profiling results to identify a single mechanism with which to construct a chemical category. However, the endpoint specific profilers identify a single structural alert that has been shown to be related to genotoxic carcinogenicity (Figure 5.3). These results enable one to be reasonable confident that the identified structural alert is related to the endpoint of interest and thus is suitable for category formation.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

1 [target]	
Structure	
<ul style="list-style-type: none"> <input checked="" type="checkbox"/> Substance Identity <input checked="" type="checkbox"/> Physical Chemical Properties <input checked="" type="checkbox"/> Environmental Fate and Transport <input checked="" type="checkbox"/> Ecotoxicological Information <input checked="" type="checkbox"/> Human Health Hazards <input type="checkbox"/> Profile <ul style="list-style-type: none"> <input type="checkbox"/> General Mechanistic <ul style="list-style-type: none"> DNA binding by OASIS v.1.1 DNA binding by OECD <input type="checkbox"/> Endpoint Specific <ul style="list-style-type: none"> Carcinogenicity (genotox and nongenotox) alerts b... Oncologic Primary Classification 	<p>Radical</p> <p>Radical >> Radical mechanism by ROS formation</p> <p>Radical >> Radical mechanism by ROS formation >> Nitroso compounds</p> <p>SN1</p> <p>SN1 >> Carbenium ion formation</p> <p>SN1 >> Carbenium ion formation >> Nitroso compounds</p> <p>SN1 >> Glutathione-induced nitrenium ion formation</p> <p>SN1 >> Glutathione-induced nitrenium ion formation >> Nitroso compounds</p> <p>SN1 >> Non-enzymatic nitroso radical and/or nirtosonium cation formation</p> <p>SN1 >> Non-enzymatic nitroso radical and/or nirtosonium cation formation >> Nitroso compounds</p> <p>SN1</p> <p>SN1 >> Carbenium Ion Formation</p> <p>SN1 >> Carbenium Ion Formation >> N-Nitroso (alkylation)</p> <p>SN2</p> <p>SN2 >> Nitrosation-SN2</p> <p>SN2 >> Nitrosation-SN2 >> Nitroso-SN2</p> <p>Alkyl and aryl N-nitroso groups (Genotox)</p> <p>Structural alert for genotoxic carcinogenicity</p> <p>Nitrosamine Type Compounds</p>

Figure 5.3: Result of primary profiling for *N*-nitrosodimethylamine using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox.

5.3 Unreliable profiling scenario

The final common profiling scenario is one in which the mechanistic and endpoint specific profilers do not allow for the identification of a single mechanism and/or structural alert related to covalent DNA binding. Profiling results such as these are of little practical use in the development of a robust chemical category suitable for data-gap filling. In cases such as these, additional research into the potential mechanism would be required in order to develop a category. Consider the example:

5.3.1 Trichlorotoluene - genotoxic carcinogenicity

Profiling trichlorotoluene for genotoxic carcinogenicity results in a variety of potential mechanisms being identified. None of these mechanisms are supported by the endpoint specific profilers with one of the profilers suggesting a non-genotoxic mechanism (Figure 5.4). The mechanistic and endpoint specific profiling results are

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

clearly contradictory and are of little use for the development of a robust chemical category.

1 [target]	
Structure	
<ul style="list-style-type: none"> <input type="checkbox"/> Substance Identity <input type="checkbox"/> Physical Chemical Properties <input type="checkbox"/> Environmental Fate and Transport <input type="checkbox"/> Ecotoxicological Information <input type="checkbox"/> Human Health Hazards <input type="checkbox"/> Profile <ul style="list-style-type: none"> <input type="checkbox"/> General Mechanistic <ul style="list-style-type: none"> DNA binding by OASIS v.1.1 DNA binding by OECD <input type="checkbox"/> Endpoint Specific <ul style="list-style-type: none"> Carcinogenicity (genotox and nongenotox) alerts b... Oncologic Primary Classification 	<p>Acylation</p> <p>Acylation >> Direct acylation involving a leaving group</p> <p>Acylation >> Direct acylation involving a leaving group >> Geminal Polyhaloalkanes</p> <p>Radical</p> <p>Radical >> Radical mechanism by ROS formation</p> <p>Radical >> Radical mechanism by ROS formation >> Geminal Polyhaloalkanes</p> <p>Schiff base formers</p> <p>Schiff base formers >> Direct acting Schiff base formers</p> <p>Schiff base formers >> Direct acting Schiff base formers >> Geminal Polyhaloalkanes</p> <p>Schiff base formers >> Multi-step Schiff base formation</p> <p>Schiff base formers >> Multi-step Schiff base formation >> Haloalkanes Containing Electron</p> <p>SN2</p> <p>SN2 >> Nucleophilic substitution at sp3 Carbon atom</p> <p>SN2 >> Nucleophilic substitution at sp3 Carbon atom >> Geminal Polyhaloalkanes</p> <p>SN2 >> Nucleophilic substitution at sp3 Carbon atom >> Haloalkanes Containing Electror</p> <p>Michael addition</p> <p>Michael addition >> P450 Mediated Activation to Quinones and Quinone-type Chemicals</p> <p>Michael addition >> P450 Mediated Activation to Quinones and Quinone-type Chemicals :</p> <hr/> <p>Alkyl halides (Nongenotox)</p> <p>Structural alert for nongenotoxic carcinogenicity</p> <hr/> <p>Not classified</p>

Figure 5.4: Result of primary profiling for trichlorotoluene using a combination of the mechanistic and endpoint specific profilers for available for genotoxic carcinogenicity in the OECD QSAR Toolbox.

5.4 General conclusions regarding the outcome of profiling strategy results

The above section outline the common examples of profiling results that one is likely to encounter when developing chemical categories for the endpoints discussed in this guidance document. The examples raise several important issues in terms of the confidence associated with profiling results and thus the subsequent category, these being:

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

5.4.1 The use of multiple mechanistic profilers

It is perhaps tempting to suggest that if both mechanistic profilers indicate the same mechanism that one should have more confidence in the profiling results. However, this is not the case as the mechanistic profilers have been developed from a range of toxicological data sources (many of which are the same for both profilers). Thus, in the case where both mechanistic profilers trigger the same alert it is likely that the underlying structural alert has been developed from the same (or similar) data. In addition, the situation where only one of the mechanistic profilers triggers an alert does not mean that the results are of lower confidence. All that can be stated in this scenario is that the target chemical contains an alert that is outside the domain of the second profiler. As discussed, the only time where one can make a decision about confidence based on the results from the two mechanistic profilers is in the situation where the results suggest multiple, competing mechanisms.

5.4.2 The use of endpoint specific profilers

The endpoint specific profilers contain structural alerts that have been shown to be associated with specific toxicological endpoints. In contrast, the mechanistic profilers have been developed from an analysis of a range of data sources. Thus, not all of the structural alerts within them have been definitively associated with every endpoint for which covalent DNA binding is the molecular initiating event. As discussed, this means that the information in the mechanistic profilers can be supplemented with the information in the endpoint specific profilers allowing one to have increased confidence in the resulting category. The applicable primary profilers for each of the three endpoints covered by this guidance document are summarised in Table 5.1.

Table 5.1: Summary of the applicable primary profilers for the mutagenicity, chromosomal mutation and genotoxic carcinogenicity available in the OECD QSAR Toolbox

Endpoint	Primary profilers (mechanistic)	Primary profilers (endpoint)
Mutagenicity	DNA binding by OECD DNA binding by OASIS v1.1	<i>In vitro</i> mutagenicity (Ames test) by ISS
Chromosomal mutation	DNA binding by OECD DNA binding by OASIS v1.1	<i>In vivo</i> mutagenicity (micronucleus) by ISS
Genotoxic carcinogenicity	DNA binding by OECD DNA binding by OASIS v1.1	Carcinogenicity (genotoxic and non-genotoxic) by ISS OncoLogic primary classification

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

6 Secondary profilers relevant to genetic toxicology and genotoxic carcinogenicity

In addition to the primary profilers, a number of secondary profilers are also of use in category formation for genetic toxicity and genotoxic carcinogenicity endpoints. These profilers are summarised in Table 6.1. In contrast to the initial battery of profilers which are used in combination with one another, the secondary profilers are best utilised individually to help sub-categorise a chemical category. Such sub-categorisation is often needed to refine the structural domain of a chemical category allowing transparent structure-activity relationships to be developed. However, it is important to state that these secondary profilers are based on various measures of chemical similarity or the presence of functional groups. Thus, they should be used cautiously in order to ensure that sub-categorisation is carried out in a logical and transparent manner. It is not the intention to delete or exclude structural analogues from a category for unknown reasons.

Table 6.1: Secondary profilers relevant to genetic toxicity and genotoxic carcinogenicity endpoints available in the OECD QSAR Toolbox

Profiler name	Type	Number of alerts
Organic functional groups	Empiric	489
Organic functional groups (nested)	Empiric	489
Organic functional groups, Norbert Haider (checkmol)	Empiric	204
Organic functional groups (US EPA)	Empiric	467
Chemical elements	Empiric	34

The most commonly utilised and useful secondary profilers are the organic functional groups and chemical elements profilers. These profilers allow the user to develop sub-categories based on the presence or absence of common organic functional groups such as carbonyl, nitro or many others. In addition, the chemical elements profiler allows sub-categories to be developed based on the presence or absence of chemical elements. A combination of one of the organic functional group profilers and the chemical elements profiler can provide useful sub-categories depending on the makeup of the chemical category. The choice of which of the three organic functional group profilers to use is largely dependent on the data within the category one wishes to sub-categorise. However, as a general approach, it is advised to use the organic functional group profiler as it relates to well established organic functional groups and thus is the most interpretable. The remaining two organic functional group profilers should be used in cases where the organic functional group profiler does not provide a satisfactory sub-category. In

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

In addition to the organic functional group profiler, the chemical elements profiler is also a useful secondary profiler. This profiler encodes the chemical elements within a molecule allowing the user to exclude a given element or sets of elements. This would become useful during the fine-tuning of a chemical category as it allows the user to restrict the category members to those whose elements are the same as are present in the target chemical.

6.1 Defining the structural domain of a chemical category

One of the key functions of the secondary profilers is in the definition of the structural domain of the chemical category. It is important that chemicals containing (significantly) different elements and functional groups are removed from the category. Typically this is achieved using a combination of the organic functional group and chemical elements profilers. Ideally, the category resulting from the primary profiling should contain only chemicals with the same elements and functional groups as the target chemical (those identified in the 'target menu' of the sub-categorisation window). However, this is not always possible and using such a tight structural domain results in the elimination of too many analogues from the category. In these instances, one can include more functional groups (by selecting them 'by hand' in the 'analogues menu' in the sub-categorisation window) to increase the number of analogues in the resulting category. A useful approach to ensure that the structural domain of the chemical category is suitable for subsequent data gap filling is as follows:

1. Profile the endpoint specific category using the organic functional group profiler removing all chemicals that contain functional groups not present in the target chemical.
2. Inspect the resulting chemical category – if it contains sufficient analogues (that one considers) suitable to fill the data gaps of interest then no further sub-categorisation is required (the absolute minimum for read across is a category containing the target chemical and a single analogue as this would allow for one-to-one read across. However, ideally one would like a category in which trend analysis and/or read across predictions could be made on a many-to-one basis. Thus, one wants a category containing at least two or three analogues if possible. For general guidance on grouping, chemical category formation and read across see [19]).
3. If steps 1 and 2 result in insufficient chemicals considered suitable for data gap filling, then re-profile the endpoint specific category using the organic functional group profiler. However, instead of removing all chemicals as before, additional simple non-ionisable organic functional groups not present

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

in the target chemical should be included (simple alkyl groups for example). This increases the likelihood that there will be sufficient chemicals included in the resulting chemical category to allow for data gap filling.

When profiling for organic functional groups 'by hand' (as in step 3 above) it is extremely important to visually inspect the types (i.e. the chemical structures and associated functional groups) of chemicals that one is eliminating. The chemicals that will be eliminated can be visualised by right clicking on the 'sub-categorisation' window and selecting 'display selected'. One approach when sub-categorising organic functional groups in this way is to try to produce a chemical category in which a suitable descriptor (for example hydrophobicity or chemical reactivity) is responsible for (the majority of) the trends in toxicity. Doing so will ensure that any subsequent predictions made by read across or trend analysis are as transparent and interpretable as possible.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

7 Initial category formation using the primary profilers to define the mechanistic domain

The formation of the initial chemical category is carried out by profiling the relevant databases for genetic toxicity and genotoxic carcinogenicity using one of the two mechanistic profilers ('DNA binding by OECD' or 'DNA binding by OASIS v1.1' profilers). The profiling and formation of the initial category is carried out as follows:

1. Profile the target chemical using the relevant primary profilers.
2. Using the results of one of the mechanistic profilers, profile the relevant databases to the endpoint of interest for chemical analogues acting via the same mechanism as the target chemical.¹
3. The resulting category is termed the 'initial category'.

It is frequently necessary to perform a sub-categorisation of the initial category using one or more of the primary profilers relevant to genetic toxicity and genotoxic carcinogenicity. This is to ensure that the chemical category relates to a single mechanism of action. Sub-categorisation of a category is carried out as follows:

1. Profile the initial category with the primary profiler that was used to develop the initial category. This profiling will identify the mechanisms (if using a mechanistic profiler) or structural alerting groups (if using an endpoint specific profiler) that are present in the initial category. These mechanisms (or structural alerting groups) are displayed in the sub-categorisation window.
2. Eliminate analogues from the initial category that contain additional mechanisms (or structural alerting groups if using an endpoint specific profiler). Ensure that the 'differ from target by' option in the sub-categorisation window is set to 'at least one category'.
3. Repeat steps 1 and 2 for the other relevant primary profilers. The order in which the primary profilers are used to perform the additional sub-categorisations is not important.

¹ Profiling the relevant databases means searching the databases for analogues.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

Specific examples will now be discussed to show how the above process works for each of the three endpoints. The following examples assume the user is familiar with the workflow of the OECD QSAR Toolbox. Thus, multiple steps and keystrokes in the workflows are omitted with only key screenshots being included. All of the profiling steps detailed should be carried out with the 'differ from target by' option set to 'at least one category' unless otherwise stated. It is important that the user is familiar with the general approach to category formation within the OECD QSAR Toolbox (those users who are not familiar should see the guidance material [1]).

7.1 Mutagenicity

This discussion relates to ethanal for the development of a chemical category suitable for data gap filling for Ames mutagenicity. The results of profiling using the 'DNA binding by OECD', 'DNA binding by OASIS v1.1', and '*in vitro* mutagenicity (Ames test) by ISS' profilers indicate that ethanal is a potential Schiff base former, as a structural alert is identified within the 'DNA binding by OECD' and '*in vitro* mutagenicity (Ames test) by ISS'. This information can be used to develop an initial category using the 'DNA binding by OECD' profiler to select category analogues from the two applicable databases available in the OECD QSAR Toolbox (Bacterial mutagenicity ISSSTY and Genotoxicity OASIS databases). Doing so results in an initial category consisting of 68 chemicals (including the target chemical).

The following sub-categorisations are then required to ensure the category contains analogues acting via a single mechanism of action:

- Sub-categorisation of the initial category of 68 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). Figure 7.1 highlights the additional electrophilic mechanisms that are present in the initial category. This sub-categorisation results in a category of 44 chemicals.
- Sub-categorisation of the category of 44 chemicals using the 'DNA binding by OASIS v1.1' profiler. This sub-categorisation results in a category of 35 chemicals.
- Sub-categorisation of the category of 35 chemicals using the '*in vitro* mutagenicity (Ames test) by ISS' profiler. This sub-categorisation results in a category of 32 chemicals.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

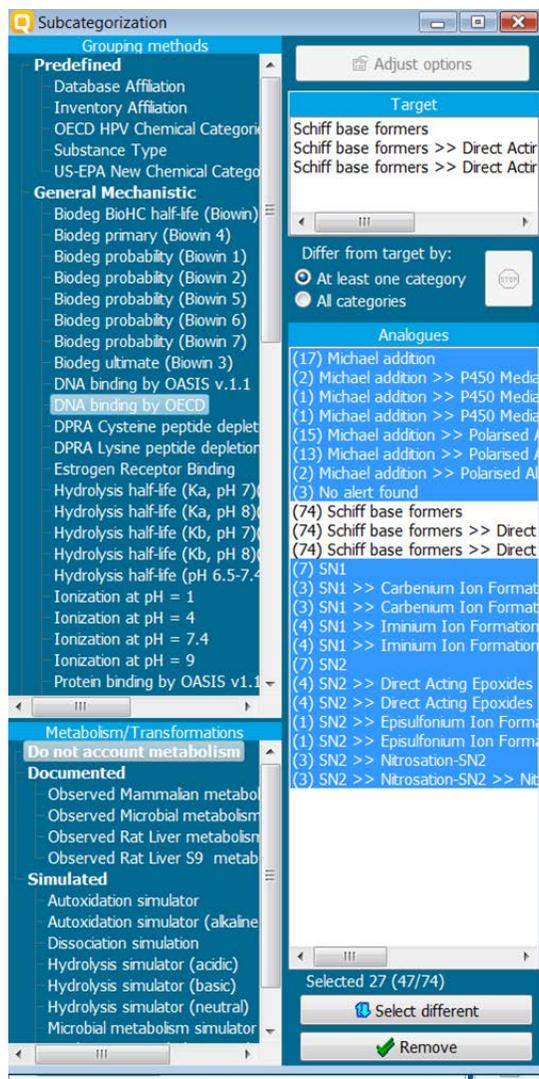


Figure 7.1: Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

7.2 Chromosomal mutation

An analogous category formation and sub-categorisation process using the primary profilers can also be carried out for chromosomal mutation using ethylene oxide as the target chemical. Three primary profilers are relevant for this endpoint (both *in vitro* and *in vivo*) these being: 'DNA binding by OECD', 'DNA binding by OASIS v1.1' and '*in vivo* mutagenicity (micronucleus) by ISS'. The three primary profilers are in agreement that the presence of the epoxide moiety in ethylene oxide is the most likely route to toxicity. This mechanistic information can be used to select analogues from the three relevant databases (genotoxicity OASIS, micronucleus ISSMIC and micronucleus OASIS) to create an initial category using the 'DNA binding by OECD' profiler. This results in an initial category of 360 chemicals (including the target chemical).

The following sub-categorisations are then required to ensure the category contains analogues acting via a single mechanism of action:

- Sub-categorisation of the initial category of 360 chemicals using the 'DNA binding by OECD' profiler (the profiler that was used to develop the initial category). Figure 7.2 highlights the additional electrophilic mechanisms that are present in the initial category. This sub-categorisation results in a category of 229 chemicals.
- Sub-categorisation of the category of 229 chemicals using the 'DNA binding by OASIS v1.1' profiler. This sub-categorisation results in a category of 192 chemicals.
- Sub-categorisation of the category of 192 chemicals using the '*in vivo* mutagenicity (micronucleus) by ISS' profiler. This sub-categorisation results in a category of 84 chemicals.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

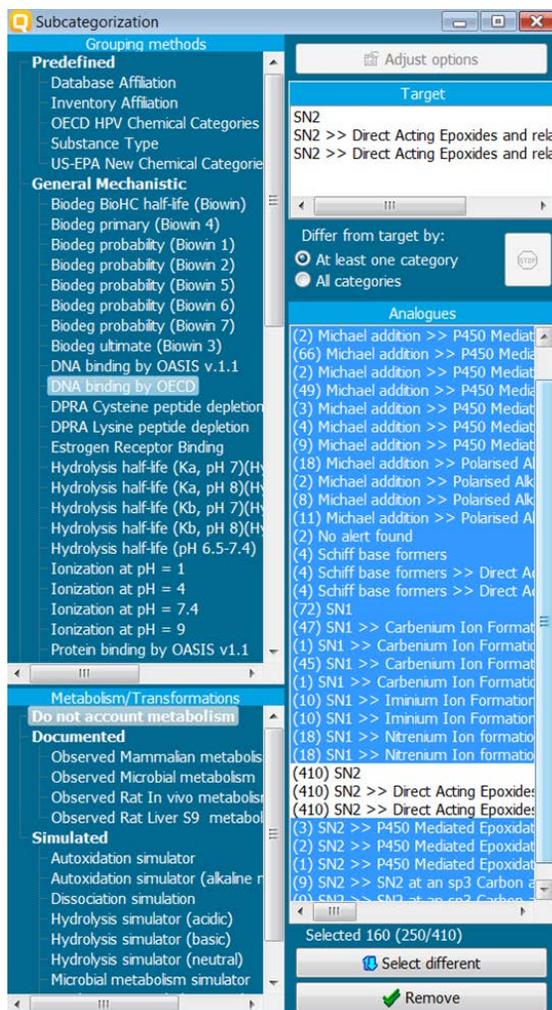


Figure 7.2: Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OECD' profiler

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

7.3 Genotoxic carcinogenicity

The formation of an initial category and then the subsequent sub-categorisation process for genotoxic carcinogenicity can be achieved using four primary profilers, these being: 'DNA binding by OECD', 'DNA binding by OASIS v1.1', 'carcinogenicity (genotoxicity/non-genotoxicity) by ISS' and 'OncoLogic primary classification'. Using N-nitrosodimethylamine to illustrate this point an initial category can be created by selecting analogues from the Carcinogenic potency database CPDB and Carcinogenicity ISSCAN database using the 'DNA binding by OASIS v1.1' profiler based on the presence of the nitrosation mechanism (and is supported by the other three profilers). An initial chemical category can be developed consisting of 149 chemicals (including the target chemical).

The following sub-categorisations are then required to ensure the category contains analogues acting via a single mechanism of action:

- Sub-categorisation of the initial category of 149 chemicals using the 'DNA binding by OASIS v1.1' profiler (the profiler that was used to develop the initial category). Figure 7.3 highlights the additional electrophilic mechanisms that are present in the initial category. This sub-categorisation results in a category of 113 chemicals.
- Sub-categorisation of the category of 113 chemicals using the 'DNA binding by OECD' profiler. This sub-categorisation results in a category of 93 chemicals.
- Sub-categorisation of the category of 93 chemicals using the 'carcinogenicity (genotoxicity/non-genotoxicity) by ISS' profiler. This sub-categorisation results in a category of 80 chemicals.
- Sub-categorisation of the category of 80 chemicals using the OncoLogic primary classification profiler. This sub-categorisation results in a category of 63 chemicals.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

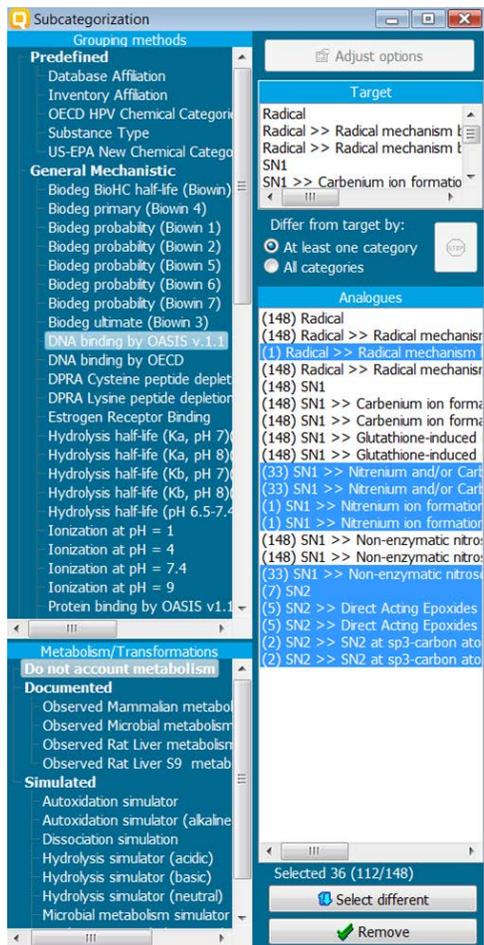


Figure 7.3: Sub-categorisation window showing the additional mechanisms identified using the 'DNA binding by OASIS' profiler

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

7.4 General conclusions regarding the sub-categorisation with the primary profilers

The examples discussed for the endpoints covered by this guidance document highlight the importance of performing a sub-categorisation with the applicable primary profilers. Such sub-categorisations are important to ensure that the resulting categories consist of chemicals acting via a single mechanism of action related to covalent DNA binding. In addition, the sub-categorisations carried out using the endpoint specific primary profilers ensure that only analogues that contain the same structural alerts present in the target chemical are included in the category. This sub-categorisation should be carried out in a step wise manner starting with the mechanistic profilers and then using the relevant endpoint specific profilers to refine the chemical category further. This will ensure a single mechanism of action is applicable for all category members, which is the main aim of this sub-categorisation process. This approach to sub-categorisation is in keeping with the OECD guidance on the grouping of chemicals [22]).

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

8 Definition of the structural domain and data-gap filling using read across

This section will outline how to use the profiling results and subsequent chemical category that was developed for ethanal in section 7.1 to fill a (hypothetical) data-gap for the TA 100 strain of *Salmonella* in the Ames mutagenicity assay. It is important that the initial profiling and sub-categorisation steps outlined in section 7.1 resulting in a category of 32 chemicals (including the target chemical) with a well-defined mechanistic domain. However, before attempting data-gap filling it is important to define the structural domain of the category using the secondary profilers.

8.1 Structural domain definition using the secondary profilers

The structural domain of the category can be defined by sub-categorising using a combination of the empiric profilers (removing all chemicals from the category that contain elements and functional groups not present in ethanal). This sub-categorisation process is analogous to that carried out using the primary profilers.

1. Sub-categorisation of the category of 32 chemicals using the chemical elements profiler. This sub-categorisation results in a category of 27 chemicals.
2. Sub-categorisation of the category of 27 chemicals using the organic functional group profiler. This sub-categorisation in a final category of 10 chemicals (including the target chemical).

8.2 Data-gap filling via read across

The sub-categorisation carried out using the primary and secondary profilers result in a category that has a well defined mechanistic (defined as a result of the sub-categorisation in section 7.1) and structural (defined in section 8.1) domains. This category of 10 chemicals (of which one is the target chemical) can now be used to fill the (hypothetical) data gap for ethanal in the TA 100 strain of *Salmonella* in the presence of the S9 liver fraction. Inspection of the toxicological data for the nine category members shows seven of them to have toxicological data. A read across prediction using the five most similar chemicals (using hydrophobicity as the measure of chemical similarity - the default approach in the OECD QSAR Toolbox) results in ethanal being predicted as negative in the TA 100 strain of *Salmonella* (Figure 8.1). This prediction is in keeping with the available toxicological data in the OECD QSAR Toolbox for ethanal.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

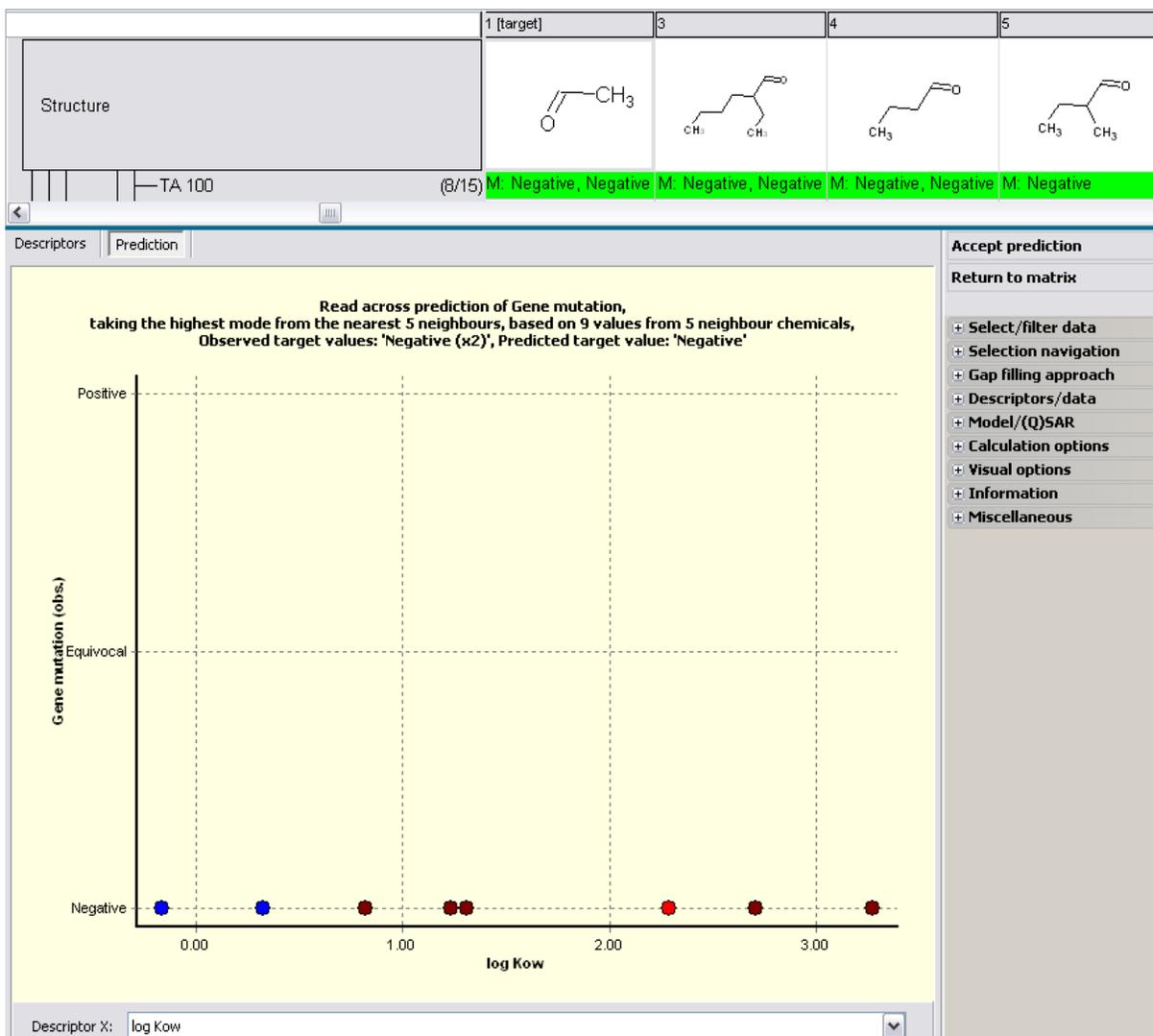


Figure 8.1: Read across prediction made for the Ames mutagenicity for ethanal in the TA 100 strain of *Salmonella* in the presence of the S9 liver fraction.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

9 General approach for the development of chemical categories for genotoxicity

The following outline can be considered a good general approach for the development of chemical categories for genetic toxicity and genotoxic carcinogenicity. These instructions are summarised in a flow chart (Figure 9.1).

1. Profile the target chemical using the two mechanistic profilers; 'DNA binding by OECD' and 'DNA binding by OASIS v1.1'.
2. Sub-categorisation using a combination of the mechanistic profilers to eliminate chemicals that contain additional potential covalent mechanism of action. Bear in mind that a robust category is applicable to a single mechanism of action.
3. Sub-categorisation of the category using an appropriate endpoint specific profiler. These profilers identify known structural alerts related to endpoint specific toxicological data. Eliminate any chemicals that contain additional structural alerts that do not occur in the target chemical. In a robust chemical category the endpoint profiling will identify a structural feature in keeping with the mechanistic profiling results.
4. Sub-categorisation using the secondary profilers in order to define the structural domain. One should use a combination of the empiric profilers (it is recommended to use the organic functional group and chemical elements profilers in the majority of cases) to restrict the structural domain of the category so that it is similar to that of the target chemical. The guiding principle should be towards the descriptor that one will use in any subsequent read across or trend analysis. This helps keep any predictions made using read across or trend analysis as transparent as possible. It is worth recalling that sometimes this profiling step requires the inclusion of analogous containing simple organic functional groups that are not present in the target chemical.
5. Always ensure that the data used in any read across or trend analysis predictions are quality checked and that unusual or outlying data within a category are investigated before use. Please remember that the OECD is not responsible for the quality of the data within the OECD QSAR Toolbox.
6. Create the appropriate reporting format in the OECD QSAR Toolbox (see guidance available from [1]).

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

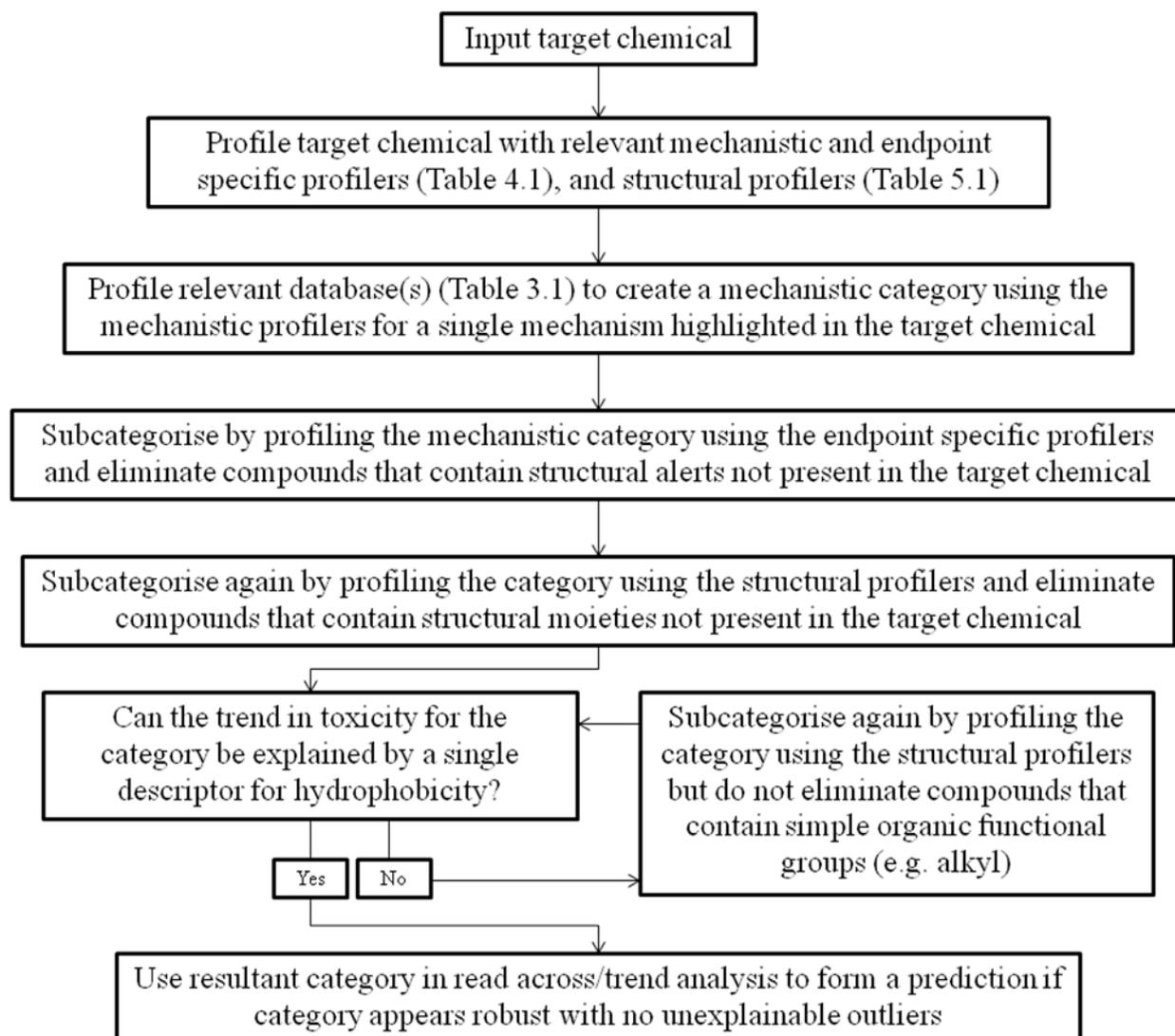


Figure 9.1: General scheme for category formation for genetic toxicity and genotoxic carcinogenicity

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

10 References

1. OECD, *The OECD (Q)SAR Application Toolbox Guidance Documents* available from http://www.oecd.org/document/54/0,3343,en_2649_34379_42923638_1_1_1_1_00.html#Guidance_Documents_and_Training_Materials_for_Using_the_Toolbox, 2010.
2. Schultz, T.W., Adverse outcome pathways: A way of linking chemical structure to in vivo toxicological hazards, in *In silico Toxicology: Principles and Applications*, M.T.D. Cronin and J.C. Madden, Editors. 2010, Royal Society of Chemistry: Cambridge, UK.
3. Benigni, R. and C. Bossa, Structure alerts for carcinogenicity, and the Salmonella assay system: A novel insight through the chemical relational databases technology. *Mutation Research*, 2008. **659**: p. 248-261.
4. Benigni, R., et al., *Development of structural alerts for the in vivo micronucleus assay in rodents*. http://ecb.jrc.ec.europa.eu/DOCUMENTS/QSAR/EUR_23844_EN.pdf, 2009.
5. Enoch, S.J. and M.T.D. Cronin, *A review of the electrophilic reaction chemistry involved in covalent DNA binding*. *Critical Reviews in Toxicology*, 2010. **40**: p. 728-748.
6. Ashby, J. and R.W. Tennant, Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutation Research*, 1988. **204**: p. 17-115.
7. Ashby, J. and R.W. Tennant, Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemical tested by the US NTP. *Mutation Research*, 1991. **257**: p. 229-306.
8. Woo, Y.-T. and D.Y. Lai, Mechanisms of action of chemical carcinogens and their role in structure-activity relationships (SAR) analysis and risk assessment, in *Quantitative Structure-Activity Relationship (QSAR) Models of Mutagens and Carcinogens*, R. Benigni, Editor 2003, CRC Press LLC.
9. Ames, B.N., B. Garry, and L.A. Herzenberg, *The genetic control of the enzymes of histidine biosynthesis in Salmonella typhimurium*. *Journal of General Microbiology*, 1960. **22**(2): p. 369-378.
10. Mortelmans, K. and E. Zeiger, *The Ames Salmonella/microsome mutagenicity assay*. *Mutation Research*, 2000. **455**: p. 29-60.

Strategies for grouping chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity

11. Kirsh-Volders, M., et al., The in vitro micronucleus test: a multi-endpoint assay to detect simultaneously mitotic delay, apoptosis, chromosome breakage, chromosome loss and non-disjunction. *Mutation Research*, 1997. **392**: p. 19-30.
12. Peto, R., et al., The TD₅₀: A proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. *Environmental Health Perspectives*, 1984. **58**: p. 1-8.
13. Aptula, A.O. and D.W. Roberts, Mechanistic applicability domains for nonanimal-based prediction of toxicological end points: General principles and application to reactive toxicity. *Chemical Research in Toxicology*, 2006. **19**(8): p. 1097-1105.
14. Mekenyan, O., et al., Identification of the structural requirements for mutagenicity by incorporating molecular flexibility and metabolic activation of chemicals I: TA100 Model. *Chemical Research in Toxicology*, 2004. **17**(6): p. 753-766.
15. Serafimova, R., et al., Identification of the structural requirements for mutagenicity, by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. *Chemical Research in Toxicology*, 2007. **20**(4): p. 662-676.
16. JRC, *ToxTree*. (available from: <http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE>), 2009.
17. Benigni, A. and Bossa, C., Mechanisms of chemical carcinogenicity and mutagenicity: A review with implications for predictive toxicology. *Chemical Reviews*, 2011. **111**: p2507- 2536
18. Woo, Y.-T., et al., An integrative approach of combining mechanistically complementary short-term predictive tests as a basis for assessing the carcinogenic potential of chemicals. *Journal of Environmental Science and Health, Part C*, 1998. **16**(2): p. 101-122.
19. OECD, *Guidance on Grouping of Chemicals*. (available from: http://www.oecd.org/LongAbstract/0,3425,en_2649_34379_39404999_119829_1_1_37407,00.html), 2007.

OECD

2, rue André Pascal
75775 Paris Cedex 16
France

Tel.: +33 1 45 24 82 00

Fax: +33 1 45 24 85 00